Neural Networks as Catalysts for Enhanced Lossless Compression in Satellite Communications

Thibault Piana, Matthieu Arzel, Abdeldjalil Aissa El Bey, Alain Thomas

Neural Networks as Catalysts for Enhanced Lossless Compression in Satellite Communications

Thibault PIANA[†], **Matthieu ARZEL**[††], **Abdeldjalil AÏSSA-EL-BEY**[††], ***and*** **Alain THOMAS**[†], ***Nonmembers***

SUMMARY Satellite communication systems enter the New Space era with a growing pressure on the ground segment, whose architecture converges to that of mobile networks with Centralized Radio Access Networks (C-RAN). Lossless data compression is a key feature for such systems, for which we propose an alternative strategy to conventional methods, such as Linear Predictive Coding (LPC), by using nonlinear predictors based on neural networks that show substantial performance improvements in high-SNR conditions (compression ratio increased by more than 10 percentage points with some architectures), while preserving signal quality. These neural networks demonstrate robustness by maintaining comparable performance with LPC in low-SNR conditions ($E_s/N_0 < 10$dB). We provide a comprehensive comparison between these methods under varying SNR conditions and delve into the influence of network parameters on performance. Our results suggest that neural networks could serve as an effective tool for improving lossless data compression in modern communication systems.

key words: *C-RAN, GSaaS, Lossless compression, Satellite communications, Neural networks, LPC*

1. Introduction

1.1 Inspiration and Challenges

Ground Station as a Service (GSaaS) was introduced by Amazon Web Services (AWS), KSAT and Microsoft Azure and has helped many companies to offer services — especially for Earth Observation (EO) — although they could not afford expensive infrastructures, as ground stations and complex terrestrial networks. Based on the experience of the cost-efficient mobile networks, they have exploited the benefits of centralized radio access networks (C-RAN) [1]. A key feature of C-RANs is their centralized nature, which allows cost-effective maintenance, efficient resource allocation, and improved operational benefits, especially in large-scale deployments. However, as illustrated in Figure 1, it requires splitting the signal processing chain so that most of or all the digital processing is done in centralized datacenters, the base/ground stations being only responsible for transmitting/receiving raw samples provided to/by their analog front-ends. Therefore, data transfers in C-RAN occur primarily through raw complex samples, representing the baseband signal, and thus require high bandwidths on the network interconnecting the base/ground stations and the datacenters. To tackle this issue, efficient and adaptive lossless compression methods are sought [1]–[3], and especially in satellite communications which use very large bandwidths such as in the Ka band for EO applications whereas in mobile communications signals are the aggregation of relatively small bandwidths. To address these challenges, new research directions are being explored in the realm of data compression, seeking high compression ratios, signal quality preservation, reduced network load, and effective handling of various waveforms.

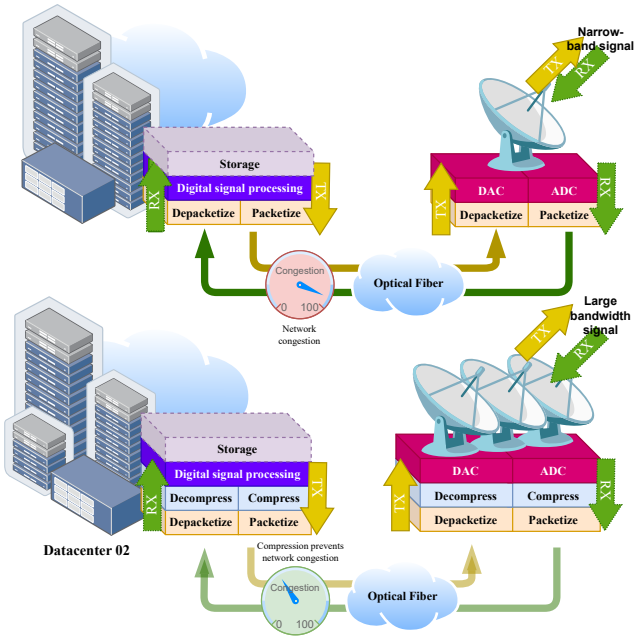


Fig. 1: Overview of remote/centralized modem processing architectures used in satellite communications systems. Compressing the digitized raw signals is the key to limit the congestion with an affordable network infrastructure.

1.2 Our Contributions

This paper primarily presents a novel exploration of a nonlinear predictive system based on neural networks for lossless data compression. This strategy enhances the data compression efficiency in high Signal-to-Noise Ratio (SNR) environments, providing an edge over the existing Linear Predictive Coding (LPC) methods in Centralized Radio Access Networks (C-RAN) of satellite communication systems. Additionally, our research advances a simple and adaptive network architecture as well as a training methodology which focuses on specific ranges of noise, resulting in improved performance in noise management and signal compression.

[††]The author is from *IMT Atlantique*, Lab-STICC, UMR CNRS 6285, F-29238, Brest

[†]The author is from *Safran Data Systems*, Space and Communication Lab, Les Ulis

Lastly, we provide a detailed comparison between the neural network-based predictor and traditional LPC, offering a comprehensive understanding of their performance under varying SNR conditions. This comparison lays the foundation for future research and improvements in data compression techniques.

2. Related works

2.1 Predictive coding

Predictive compression is an effective data reduction technique that leverages historical data patterns to predict future values. In essence, it works by creating a model of the data being compressed, and then using this model to predict subsequent data points. The actual data are then replaced with the difference between the real and predicted values, commonly referred to as the residual or error. Since the residuals tend to be smaller numbers, the entropy of the residual is also smaller, leading to efficient compression.

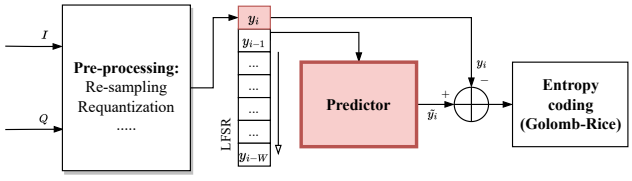


Fig. 2: Classical architecture for compression/decompression scheme using predictive coding.

2.2 Linear Predictive Coding (LPC) for Data Compression

Linear Predictive Coding (LPC) plays a pivotal role in predictive coding because of its efficiency. Essentially, LPC predicts future data points by deploying a linear combination of past samples. It can be considered as a state-of-the-art method due to its performance and its usage is various fields, including digital communications [2]–[7]. This method is also very popular in audio codecs, such as FLAC or SHORTEN, for its robust performance [8], [9]. Although LPC offers advantages such as efficient coefficient computation and easy hardware implementation, it faces difficulties when the number of samples per symbol is reduced by resampling, as shown in Figure 3. Reducing the number of samples per symbol can result in predictions that are not as accurate. This can compromise the efficiency of compression. This is because the remaining residual signal may not follow a typical Gaussian distribution, which adversely affects the performance of the Golomb-Rice coder. To increase the compression ratio, it is then necessary to conceive better predictors which can also handle low oversampling. Linear alternatives exist, such as ARMA decomposition, but struggle with their own limitations, such as parallelization challenges, stability issues, and minor performance improvements, which restrict their application in data compression.

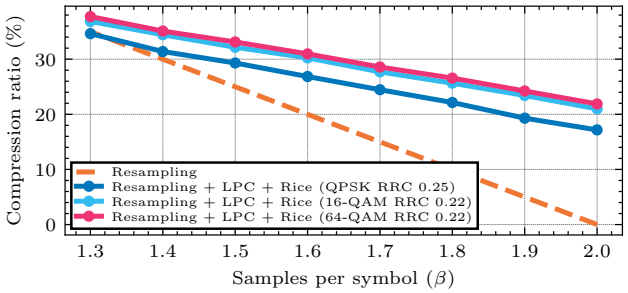


Fig. 3: Graph depicting the progression of the overall compression ratio for several waveforms. Signals are processed using LPC-based compression after resampling. Original signals are sampled with 2 samples per symbol.

2.3 Neural Predictors for Data Compression

Under high SNR, the signals used in satellite communication exhibit characteristics such as stationarity or cyclostationarity, constant power over time, known envelope, and a constant dynamic range. This regularity allows us to consider the application of a nonlinear predictive system. Despite their implementation challenges and complexity, nonlinear systems offer higher sensitivity than linear systems, which makes them suitable for stable datasets. Neural network (NN) predictors have already found applications in various fields, as their nonlinear nature can enable superior predictions. In the medical domain, they have been used for the efficient compression of EEG signals [10]. Similarly, they have been used to compress large amounts of satellite telemetry data before transmission [11]. Despite these varied applications, it is worth noting that such methods have not been explored yet for use in the field of telecommunication and satellite telecommunications. The effectiveness of neural network predictors in these varied applications, however, highlights their considerable potential in signal compression.

Our objective is then to study how efficient a nonlinear predictor based on NN can be and how the addition of noise can affect the compression performance.

3. Experimental Setup

The present study utilized a set of computer-generated signals, synthesized via Matlab. These signals span across various waveforms pertinent to distinct sectors of wireless communications, including but not limited to terrestrial, aeronautical and satellite communications, thereby fortifying the real-world applicability of our methodology. The datasets are made of sampled baseband I/Q signals. Before feeding the network, signals are resampled at a predefined sampling rate, quantized on 12 bits, and normalized based on the signal level. Subsequently, the signal was divided into fragments according to a sliding window of size W.

Network training is carried out offline, separately from their operational deployment. This approach confers several advantages. The resultant network weights, derived from this offline training, are pre-shared between the compressor

and decompressor ends before any data transmission. Importantly, this eliminates the need for weight transmission during compression, enhancing the efficiency and security of the compression and decompression process. This strategy ensures the operational readiness of the networks while minimizing computational overhead during their active use in data compression and decompression. Figure 4 presents the neural network used in our study. It is a multi-layer feed-forward network with four layers. The chosen architecture is relatively simple, yet powerful, with careful consideration given to maintaining a balance between complexity and performance. We also designed it to be really adaptable to the resources available.

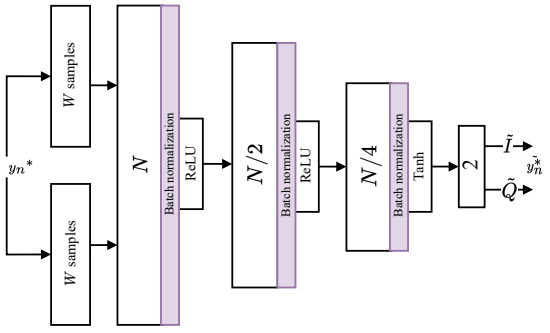


Fig. 4: Architecture of our prediction network

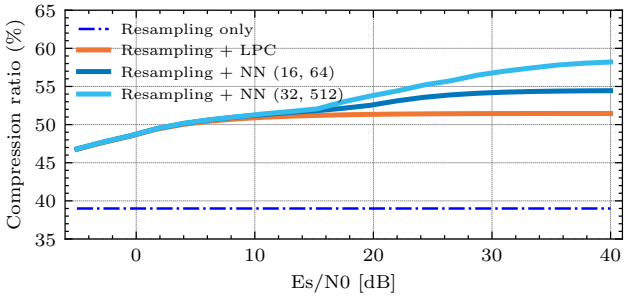
The usage of Batch Normalization (BatchNorm) is highly beneficial for noisy data. By normalizing inputs across each mini-batch, it stabilizes learning, reduces training epochs required, and enhances robustness. It makes the neural network less sensitive to noise, mitigating issues such as vanishing or exploding gradients. Thus, simulations showed that it improves the stability and accuracy of predictions and the compression performance in noisy conditions.

The architecture was designed with a deliberate intent to keep the model 'light', using a limited number of layers and neurons. The intent was to enable easy hardware implementation and parallelization while delivering strong performance. The network was trained on various sets of baseband I/Q signals, striving to predict the most likely next sequence in the data, thus reducing the residual error and improving the compression ratio.

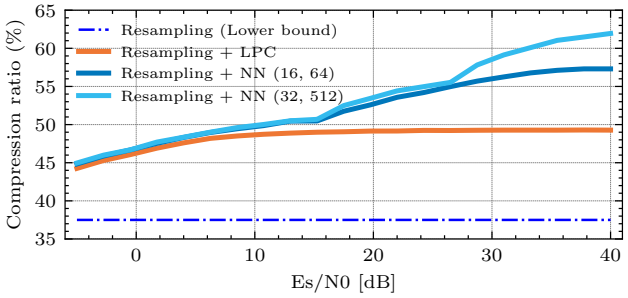
In our study, we trained the networks so that the parameters are learned specifically for successive $E_S/N0$ ranges of 7.5dB. Our simulations show that a network trained across a wide noise range tends to underperform compared to one that is tuned to a specific noise range. In practice, this is not a significant hurdle. The noise range in which a signal resides is generally known, even if the exact signal-to-noise ratio is uncertain. This knowledge aids in effectively assigning the appropriate trained network. Moreover, the simplicity of training and implementing networks and the ease of switching between sets of parameters, facilitated by indicating the set used in a header, contribute to the feasibility and efficiency of this approach in real-world applications when multiple parameter sets are needed. However, this training approach can potentially induce threshold effects at the borders of various noise ranges, leading to the observed "wave-like" patterns in the curves in Figure 5.

If the following examples, we took (W, N) equal to $(16, 64)$ and $(32, 512)$, which correspond to two architectures of 4,900 and 180,000 parameters.

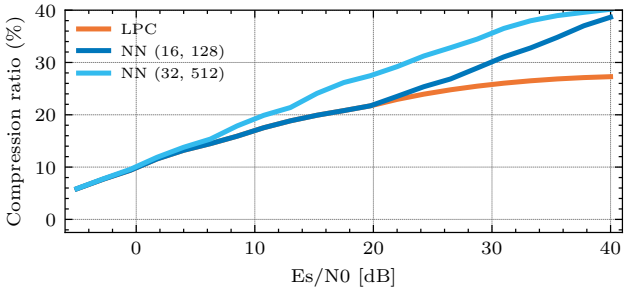
4. Results and Discussion



(a) 16-QAM signal shaped with RRC of 0.22 rolloff. Resampling is 1.25 sample per symbol



(b) 8-PSK signal shaped with RRC of 0.25 rolloff. Resampling is 1.22 sample per symbol



(c) GMSK signal with $BT = 0.3$

Fig. 5: Compression ratio obtained for various signals and SNRs. All the signals have an original sampling rate of 2 samples per symbol. Simulations showed that performing resampling before applying lossless compression increases global compression ratio. NNs were specifically trained for successive $E_S/N0$ ranges of 7.5dB. The parameter mentioned for each network is the tuple (W, N).

Our experiments show a clear advantage for the NN-based compression systems over LPC in high-SNR environments ($E_S/N0 > 10$ dB). Compared to LPC, our NN solutions offer compression ratios increased by up to 12

percentage points. Such advantage is primarily displayed in scenarios requiring high-accuracy data transmission and environments characterized by low noise. The NN-based system is proving to be a valuable asset in these contexts.

In contrast, under low-SNR conditions, the performance of the NN-based system converges to that of LPC-based. Note that LPC can be seen as a really simple case of neural network since training a network with no activation function (*i.e.*, without non-linearities) will actually find the LPC coefficients [12]. The inherent randomness of noise introduces some unpredictability that challenges the predictive abilities of the neural network. The network's performance is thus reduced to a level similar to that of LPC in these noisy situations. This particular behavior, nonetheless, showcases the system's resilience, as it continues to maintain reliable performance even in highly noisy conditions.

Furthermore, an interesting trend was observed concerning the performance of neural networks in relation to the number of their parameters. Our results demonstrate that even with a comparatively low number of parameters (4,900), the neural networks still deliver a commendable performance. Switching from NN(16,64) to NN(32,512) allows increasing the compression ratio by up to 5 percentage points but it also requires increasing the number of parameters from 4,900 to 180,000 . Pruning NN(32,512) would certainly reduce the number of parameters without significant loss in performance. Such strategies offer an interesting perspective for balancing performance and computational efficiency.

Finally, in practical use cases, such as those using the DVB-S2 standard, our method would be particularly beneficial. Measured noise levels [13] indicate that the received signal SNR in clear weather can sometimes exceed 35 dB of $E_S/N0$, and can even approach 50 dB for high elevations (near 90°). These values are significantly higher than the operating values for all DVB-S2 modulation and coding states. In the case of using Constant Coding Modulation (CCM) or Variable Coding Modulation (VCM), there are time intervals where the use of neural predictors would be very useful, as they would be operating at their full potential. In Automatic Coding Modulation (ACM), the measured margin is so high that even with the highest modulation, it remains large enough for effective compression. In less clear weather conditions, our method could also benefit from the natural margins coming from the link budget of the line. This advocates using our neural network-based system in real-world applications, where it can leverage its strengths in high SNR environments to enhance satellite signal compression and thus the ground segment efficiency.

5. Conclusion

This paper delves into the complexities of lossless data compression in Centralized Radio Access Networks (C-RAN) for satellite communication systems, highlighting the limitations of conventional methods like Linear Predictive Coding (LPC). It presents neural network-based predictors as a promising alternative for enhancing compression performance, showing substantial improvements in high SNR scenarios while retaining data integrity. The results show that when a sufficient noise margin exists, the usage of a neural predictor is always profitable (compression ratio increased by up to 12 percentage points), and in the worst case, it has similar performance to that of a linear predictor. Since neural networks echo the performance of traditional LPC under low SNR conditions, it points towards a possible need for exploring lossy compression methods for further compression gains, and thus offering a better usage of the limited bandwidth of the ground segment network.

References

[1] J. Bartelt, P. Rost, D. Wubben, J. Lessmann, B. Melis, and G. Fettweis, "Fronthaul and backhaul requirements of flexibly centralized radio access networks," IEEE Wireless Communications, vol.22, no.5, pp.105–111, Oct. 2015.

[2] S. Nanba and A. Agata, "A new IQ data compression scheme for front-haul link in Centralized RAN," 2013 IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC Workshops), London, United Kingdom, pp.210–214, IEEE, Sept. 2013.

[3] L. Ramalho, M.N. Fonseca, A. Klautau, C. Lu, M. Berg, E. Trojer, and S. Host, "An LPC-Based Fronthaul Compression Scheme," IEEE Communications Letters, vol.21, no.2, pp.318–321, Feb. 2017.

[4] G. Einarsson, "An improved implementation of predictive coding compression," IEEE Transactions on Communications, vol.39, no.2, pp.169–171, Feb. 1991.

[5] L. Ramalho, I. Freire, C. Lu, M. Berg, and A. Klautau, "Improved LPC-Based Fronthaul Compression With High Rate Adaptation Resolution," IEEE Communications Letters, vol.22, no.3, pp.458–461, March 2018.

[6] L. Ramalho, J. Fortuna, C. Lu, M. Berg, I. Almeida, and A. Klautau, "Rate Control for Entropy-Coded LPC: Application to Packet-Based Fronthauling," IEEE Transactions on Communications, pp.1–1, 2022.

[7] G. Chen, F. Yang, K. Niu, and C. Dong, "A linear predictive coding based compression algorithm for fronthaul link in C-RAN," 2017 IEEE/CIC International Conference on Communications in China (ICCC), Qingdao, China, pp.1–6, IEEE, Oct. 2017.

[8] T. Robinson, "SHORTEN: Simple lossless and near-lossless waveform compression," Technical Report CUED/F-INFENG/TR.156, Cambridge University Engineering Department, Dec. 1994.

[9] M. Van Beurden, "Lossless audio codec comparison : Revision 4," Tech. Rep. 4, Jan. 2015.

[10] N. Sriraam, R. Kannan, and C. Eswaran, "Lossless compression of EEG data using neural network predictors," Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02., Singapore, pp.2046–2048, Nanyang Technol. Univ, Nov. 2002.

[11] A.F. Shehab, M.A. Elshafey, and T.A. Mahmoud, "Recurrent Neural Network Based Prediction to Enhance Satellite Telemetry Compression," 2020 IEEE Aerospace Conference, Big Sky, MT, USA, pp.1–11, IEEE, March 2020.

[12] I. Ibrahim and M. El-Adawy, "Determination of the LPC coefficients using neural networks," Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, WA, Australia, pp.2669–2673, IEEE, 1995.

[13] Z. Wang, F. Lu, D. Wang, X. Zhang, J. Li, and J. Li, "A Transmission Efficiency Evaluation Method of Adaptive Coding Modulation for Ka-Band Data-Transmission of LEO EO Satellites," Sensors, vol.22, no.14, p.5423, July 2022.